WHAT IS CLAIMED IS:

1. A method of making a recombinant nucleic acid, the method comprising:

providing a plurality of parental character strings corresponding to a plurality of nucleic acids, which character strings, when aligned for maximum identity, comprise at least one region of heterology;

aligning the character strings;

defining a set of character string subsequences, which set of subsequences comprises subsequences of at least two of the plurality of parental character strings;

providing a set of oligonucleotides corresponding to the set of character string subsequences;

annealing the set of oligonucleotides; and,

elongating one or more members of the set of oligonucleotides with a polymerase, or ligating at least two members of the set of oligonucleotides with a ligase, thereby producing one or more recombinant nucleic acid.

2. The method of claim 1, wherein the character strings, when aligned for maximum identity comprise at least one region of similarity.

3. The method of claim 1, wherein at least one of the parental character strings is an evolutionary or artificial intermediate.

4. The method of claim 1, wherein at least one of the parental character strings corresponds to a designed nucleic acid.

5. The method of claim 4, wherein the designed nucleic acid represents an energy minimized design for an encoded polypeptide.

6. The method of claim 1, further comprising applying one or more genetic operator to one or more of the parental character strings, or to one or more of the character string subsequences, wherein the genetic operator is selected from: a mutation of the one or more parental character strings or one or more character string subsequences, a multiplication of the one or more parental character strings or one or more character string subsequences, a fragmentation of the one or more parental character strings or one or more character string

81

subsequences, a crossover between any of the one or more parental character strings or one or more character string subsequences or an additional character string, a ligation of the one or more parental character strings or one or more character string subsequences, an elitism calculation, a calculation of sequence homology or sequence similarity of aligned strings, a recursive use of one or more genetic operator for evolution of character strings, application of a randomness operator to the one or more parental character strings or the one or more character string subsequences, a deletion mutation of the one or more parental character strings or one or more character string subsequences, an insertion mutation into the one or more parental character strings or one or more of character string subsequences, subtraction of the of the one or more parental character strings or one or more character string subsequences with an inactive sequence, selection of the of the one or more parental character strings or one or more character string subsequences with an active sequence, and death of the one or more parental character strings or one or more of character string subsequences.

7.   The method of claim 1, further comprising selecting a diplomat sequence, which diplomat sequence comprises an intermediate level of sequence similarity between two or more of the plurality of character strings.

8.   The method of claim 1, wherein the set of oligonucleotides comprise a plurality of overlapping oligonucleotides.

9.   The method of claim 1, wherein the set of character string subsequences is defined by selecting a length for the character string and subdividing at least two of the plurality of parental character strings into segments of the selected length.

10.   The method of claim 1, wherein aligning the character strings is performed in a digital computer or in a web-based system.

11.   The method of claim 1, further comprising synthesizing a set of single-stranded oligonucleotides which correspond to the set of character string subsequences, thereby providing the set of oligonucleotides.

12.   The method of claim 1, further comprising:
pooling all or part of the set of oligonucleotides;
hybridizing the resulting pooled oligonucleotides; and,

extending a plurality of the resulting hybridized oligonucleotides, wherein at least one of the resulting extended double stranded nucleic acids comprises sequences from at least two of the plurality of parental character strings.

13. The method of claim 11, further comprising denaturing the double stranded nucleic acids, thereby producing a heterogeneous mixture of single-stranded nucleic acids.

14. The method of claim 11, further comprising:

(i) denaturing the double stranded nucleic acids, thereby producing a heterogeneous mixture of single-stranded nucleic acids;

(ii) re-hybridizing the heterogeneous mixture of single-stranded nucleic acids; and

(iii) extending the resulting rehybridized double stranded nucleic acids with a polymerase.

15. The method of claim 13, further comprising repeating steps (i) (ii) and (iii) at least twice.

16. The method of claim 1, further comprising selecting the one or more recombinant nucleic acid for a desired property.

17. The method of claim 1 wherein the set of oligonucleotides is provided by synthesizing the oligonucleotides to comprise one or more modified parental character string subsequence, which subsequence comprises one or more of:

a parental character string subsequence modified by one or more replacement of one or more character of the parental character string subsequence with one or more different character;

a parental character string subsequence modified by one or more deletion or insertion of one or more characters of the parental character string subsequence;

a parental character string subsequence modified by inclusion of a degenerate sequence character at one or more randomly or non-randomly selected positions;

a parental character string subsequence modified by inclusion of a character string from a different character string from a second parental character string subsequence at one or more position;

a parental character string subsequence which is biased based upon its frequency in a selected library of nucleic acids; and,

a parental character string subsequence which comprises one or more sequence motif, which sequence motif is artificially included in the subsequence.

18. The method of claim 17, wherein the sequence motif comprises an N-linked glycosylation sequence, an O-linked glycosylation sequence, a protease sensitive sequence, a collagenase sensitive sequence, a Rho-dependent transcriptional termination sequence, an RNA secondary structure sequence that affects the efficiency of transcription, an RNA secondary structure sequence that affects the efficiency of translation, a transcriptional enhancer sequence, a transcriptional promoter sequence, or a transcriptional silencing sequence.

19. The method of claim 1, wherein the oligonucleotide set contains one or more altered or degenerate positions as compared to the corresponding subsequence of one or more parental character string.

20. The method of claim 1, further comprising selecting the one or more recombinant nucleic acid based upon its hybridization to a selected nucleic acid or to a set of selected nucleic acids.

21. The method of claim 1, wherein the one or more parental character string comprises at least two parental character strings, wherein the oligonucleotide set comprises at least one oligonucleotide member comprising a chimeric nucleic acid sequence, the at least one oligonucleotide member comprising at least two oligonucleotide member subsequences, wherein the at least two oligonucleotide member subsequences correspond to at least two subsequences from the at least two parental character strings, the at least two oligonucleotide member subsequences being separated by a crossover point.

22. The method of claim 21, wherein the crossover point is selected by identifying a plurality of parental character substrings from a plurality of the at least two parental character strings, aligning the substrings to display pairwise identity between the substrings, and selecting a point within the aligned sequence as the crossover point.

23. The method of claim 21, wherein the crossover point is selected randomly.

84

24. The method of claim 21, wherein the crossover point is selected non randomly.

25. The method of claim 21, wherein the crossover point is selected non randomly by selecting a crossover point approximately in the middle of one or more identified pairwise identity region.

26. The method of claim 21, wherein at least one crossover point for at least one oligonucleotide member is selected from a region outside of an identified pairwise homology region.

27. The method of claim 1, further comprising adding one or more oligonucleotide member of the set of oligonucleotides at a concentration which is higher than at least one or more additional oligonucleotide member of the set of oligonucleotides.

28. The method of claim 1, further comprising incubating one or more member of the oligonucleotide set with the recombinant nucleic acid and a polymerase.

29. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid from the oligonucleotide set.

30. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid produced by cleavage of a parental nucleic acid encoded by the at least one parental character string.

31. The method of claim 1, further comprising denaturing the recombinant nucleic acid, and contacting the recombinant nucleic acid with at least one additional nucleic acid produced by cleavage of a parental nucleic acid encoded by the at least one parental character string, which parental nucleic acid is cleaved by one or more of: chemical cleavage, cleavage with a DNAse and cleavage with a restriction endonuclease.

32. The method of claim 1, wherein the parental character string encodes one or more nucleic acid corresponding to one or more of protein or gene selected from: EPO, insulin, a peptide hormone, a cytokine, epidermal growth factor, fibroblast growth factor,

hepatocyte growth factor, insulin-like growth factor, an interferon, an interleukins, a keratinocyte growth factor, a leukemia inhibitory factor, oncostatin M, PD-ECSF, PDGF, pleiotropin, SCF, c-kit ligand, VEGEF, G-CSF, an oncogene, a tumor suppressor, a steroid hormone receptor, a plant hormone, a disease resistance gene, an herbicide resistance gene, a bacterial gene, a monooxygenases, a protease, a nuclease, and a lipase.

33. The method of claim 1, wherein the set of oligonucleotides comprises one or more oligonucleotide member between about 20 and about 60 nucleotides in length.

34. The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid.

35. The method of claim 34, further comprising recombining the selected recombinant nucleic acid with one or more of: a homolgous nucleic acid, and an oligonucleotide member from the set of oligonucleotides.

36. The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid, wherein the desired trait or property is selected in an in vivo selection assay or a parallel solid phase assay.

37. The method of claim 1, further comprising selecting the recombinant nucleic acid for a desired trait or property, thereby providing a selected recombinant nucleic acid, wherein the desired trait or property is selected in an in vitro selection assay.

38. The method of claim 1, further comprising deconvolution of the recombinant nucleic acid.

39. The method of claim 1, further comprising sequencing or cloning the recombinant nucleic acid.

40. The method of claim 1, wherein the recombinant nucleic acid is synthesized in vitro by assembly PCR.

41. The method of claim 1, wherein the recombinant nucleic acid is synthesized in vitro by error-prone assembly PCR.

42. The method of claim 1, wherein the parental character strings, or oligonucleotide sets are selected in a computer.

43. A method of making character strings, the method comprising:

a) providing a parental character string encoding a polynucleotide or polypeptide;

5   b) providing a set of oligonucleotide character strings of a pre-selected length that encode a plurality of single-stranded oligonucleotide sequences comprising sequence fragments of the parental character string, and complement thereof;

c) creating a set of derivatives of the parental sequence comprising sequence variant strings, the set comprising a plurality of mutations, having one mutation per variant string.

10   **44.** The method of claim 43, wherein a plurality of the plurality of single-stranded oligonucleotide sequences are overlapping in sequence.

45. The method of claim 43, further comprising applying one or more genetic operator to the parental character string, or to one or more of the oligonucleotide character strings, wherein the genetic operator is selected from: a mutation of the parental character

15   string, or one or more of the oligonucleotide character strings, a multiplication of the parental character string, or one or more of the oligonucleotide character strings, a fragmentation of the parental character string, or one or more of the oligonucleotide character strings, a crossover between any of the parental character string or one or more of the oligonucleotide character strings, or an additional character string, a ligation of the of the parental character string, or one

20   or more of the oligonucleotide character strings, an elitism calculation, a calculation of sequence homology or sequence similarity of an alignment comprising the parental character string, or one or more of the oligonucleotide character strings, a recursive use of one or more genetic operator for evolution of character strings, application of a randomness operator to the parental character string, or to one or more of the oligonucleotide character strings, a deletion

25   mutation of the parental character string, or one or more of the oligonucleotide character strings, an insertion mutation into the parental character string, or one or more of the oligonucleotide character strings, subtraction of the of the parental character string, or one or more of the oligonucleotide character strings with an inactive sequence, selection of the of the parental character string, or one or more of the oligonucleotide character strings with an active

30   sequence, and death of the parental character string, or one or more of the oligonucleotide character strings.

46. The method of claim 43, further comprising:

d) providing a set of overlapping character strings of a pre-defined length that encode both strands of the parental character string sequence; and,

e) synthesizing sets of single-stranded oligonucleotides according to the step (c) and (d).

47. The method of claim 46, further comprising:

f) assembling a library of recombinant nucleic acids by assembly PCR from the single-stranded oligonucleotides.

48. A library made by the method of claim 47.

49. The method of claim 47, further comprising:

g) selecting or screening the library for one or more recombinant polynucleotide having a desired property.

50. The method of claim 48, further comprising:

h) deconvoluting the sequence of the one or more selected polynucleotide.

51. The method of claim 46, wherein the sequence of the one or more selected polynucleotide is deconvoluted by sequencing the selected polynucleotide, or by digesting the one or more selected polynucleotide.

52. The method of claim 46, wherein the sequence is deconvoluted by positional deconvolution of the one or more selected polynucleotide.

53. The method of claim 46, further comprising reiterative shuffling or selection of the library of recombinant nucleic acids.

54. A method of facilitating recombination between two or more divergent nucleic acids, the method comprising:

aligning parental character strings corresponding to the divergent nucleic acids, thereby identifying regions of sequence identity and regions of sequence diversity;

defining a diplomat character string which is intermediate in sequence between the parental character strings;

synthesizing at least a portion of the diplomat sequence to produce a diplomat nucleic acid; and,

recombining a mixture of selected nucleic acids comprising the parental nucleic acids, or fragments thereof, and the diplomat nucleic acid.

5   **55.** The method of claim 54, wherein the diplomat nucleic acid is synthesized by synthesizing a plurality of overlapping oligonucleotides corresponding in sequence to the diplomat sequence, hybridizing the overlapping oligonucleotides, and incubating the overlapping oligonucleotides with a polymerase.

56. The method of claim 54, further comprising synthesizing a pool of
10   oligonucleotides corresponding to one or more of the parental character strings, which pool of oligonucleotides is present in the mixture of selected nucleic acids.

57. The mixture of selected nucleic acids produced by the method of claim 56.

58. A method of generating and recombining nucleic acids, the method comprising:

15   inputting a plurality of amino acid sequence character strings into a digital system;

reverse translating the amino acid character strings in the digital system into a plurality of nucleic acid character strings, wherein reverse translated nucleic acid sequences are selected for one or more of: species codon bias in a selected expression host, and optimized sequence similarity between the plurality of nucleic acid character strings; and,

20   synthesizing one or more sets of oligonucleotides corresponding to one or more reverse translated nucleic acid sequences.

59. The method of claim 58, further comprising hybridizing members of the one or more oligonucleotide sets to each other, or to a set of fragmented nucleic acids which encodes one or more amino acid polymer corresponding to one or more of the amino acid
25   sequence character strings.

60. The method of claim 59, further comprising elongating one or more resulting hybridized nucleic acids with a polymerase.

61. The method of claim 58, further comprising fragmenting one or more resulting elongated nucleic acids and hybridizing the resulting secondary fragmented nucleic

acids with each other or with members of the one or more oligonucleotide sets, or with a set of primary fragmented nucleic acids which encodes one or more amino acid polymer corresponding to one or more of the amino acid sequence character strings.

62. A method of optimizing activity of a nucleic acid, the method comprising:

5   parameterizing a set of nucleic acids or proteins to provide a set of multidimensional datapoints;

extrapolating one or more postulated multidimensional datapoint from the set of multidimensional datapoints; and,

converting the postulated multidimensional datapoint to a new character string

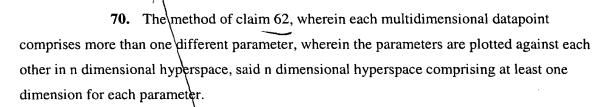10   corresponding to a postulated nucleic acid nucleic acid or protein.

63. The method of claim 62, comprising synthesizing the postulated nucleic acid or protein.

64. The method of claim 62, further comprising principle component analysis of the set of multidimensional datapoints.

15   65. The method of claim 62, comprising shuffling the postulated nucleic acid, or a subsequence thereof, with an additional nucleic acid.

66. The method of claim 62, wherein the set of nucleic acids or proteins is parameterized by correlating each residue of the nucleic acid or protein to a matrix of numeric indicators.

20   67. The method of claim 66, wherein the matrix is graphically represented as a tetrahedron, having an assigned origin at the center of the tetrahedron, with each corner represented as a numeric representation, with each residue of a nucleic acid being positioned at a different corner, thereby producing the matrix of numeric indicators.

68. The method of claim 62, comprising correlating each multidimensional

25   datapoint with an output vector to identify a relationship between a matrix of dependent Y variables and a matrix of predictor X variables.

69. The method of claim 68, wherein the correlation is performed by partial least square projections to latent structures analysis.

90

70. The method of claim 62, wherein each multidimensional datapoint comprises more than one different parameter, wherein the parameters are plotted against each other in n dimensional hyperspace, said n dimensional hyperspace comprising at least one dimension for each parameter.

5

71. A method of providing a library of recombinant nucleic acids which is enriched for a sequence of interest and selecting the library, the method comprising:

producing an initial library of at least about 10^6 recombinant nucleic acids, which initial library of recombinant nucleic acids comprises at least about 10^5 different member types, which 10^5 different member types are non-identical;

10

hybridizing the library to one or more population of nucleic acids, which one or more population of nucleic acids correspond to one or more subsequences in the different library members;

isolating members of the library which hybridize to the one or more populations of nucleic acids, thereby enriching the library of nucleic acid for members which hybridized to

15 the one or more population of nucleic acids; and,

selecting members of the resulting enriched library for one or more property of interest.

72. The method of claim 71, wherein the initial library has between about 10^9 and 10^{12} members.

73. The method of claim 71, wherein the one or more population of nucleic

20 acids is fixed to a solid substrate.

74. The method of claim 73, wherein the solid substrate comprises one or more of: a column matrix material and a nucleic acid chip.

75. The method of claim 71, wherein the initial library is produced by recombining one or more homologous nucleic acids.

25

76. The enriched library produced by the method of claim 71.

77. The method of claim 71, wherein the initial library is produced by:

providing a plurality of parental character strings corresponding to a plurality of nucleic acids, which character strings, when aligned for maximum identity, comprise at least one region of similarity and at least one region of heterology;

aligning the character strings;

defining a set of character string subsequences, which set of subsequences comprises subsequences of at least two of the plurality of parental character strings;

providing a set of oligonucleotides corresponding to the set of character string

5   subsequences;

annealing the set of oligonucleotides; and,

elongating one or more member of the set of oligonucleotides with a polymerase, thereby producing the initial library of nucleic acids.

78.   A method of generating a library of biological polymers, the method

10   comprising:

generating a diverse population of character strings in a computer, which character strings are generated by alteration of pre-existing character strings; and,

synthesizing the diverse population of character strings, which diverse population comprises the library of biological polymers.

15   **79.**   The method of claim 78, wherein the alteration comprises recombination of the pre-existing character strings.

80.   The method of claim 78, wherein the biological polymers are selected from nucleic acids, polypeptides and peptide nucleic acids.

81.   The method of claim 78, further comprising selecting members of the

20   library of biological polymers for one or more activity.

82.   The method of claim 81, further comprising filtering an additional library or an additional set of character strings by subtracting the additional library or the additional set of character strings with members of the library of biological polymers which display activity below a desired threshold.

25   **83.**   The method of claim 81, further comprising filtering an additional library or an additional set of character strings by biasing the additional library or the additional set of character strings with members of the library of biological polymers which display activity above a desired threshold.

84. An integrated system comprising a computer having a first data set comprising a first character string, a second data set comprising a second character string, software for aligning the first and second character strings, software for performing a genetic operation on the first or second character string, an output file comprising a third data set

5   comprising a third character string, the third character string comprising character string subsequences from the first and second character strings, and an oligonucleotide sequence output file comprising a plurality of overlapping oligonucleotide sequences corresponding to the third character string.

85. The integrated system of claim 84, the system further comprising an

10   oligonucleotide synthesis machine for synthesizing the plurality of overlapping oligonucleotides.

86. The integrated system of claim 84, further comprising a plurality of oligonucleotides encoded by the plurality of overlapping oligonucleotide sequences, which oligonucleotides, when incubated in one or more cycles of chain extension, produce a third

15   nucleic acid encoded by the third character string.

87. The integrated system of claim 84, wherein the system further comprises a program with an instruction set for applying one or more genetic operator to the first or second character string, or to any other character string.

88. The integrated system of claim 84, wherein the system further comprises a

20   program with an instruction set for applying one or more genetic operator to the first or second character string, or to any other character string, wherein the genetic operator is selected from: a mutation, a multiplication, a fragmentation of the string or strings, a crossover between one or more strings, a ligation of strings, an elitism calculation, an alignment, a calculation of sequence homology or sequence similarity, a recursive use of one or more genetic operator for

25   evolution of character strings, randomness, a deletion mutation, an insertion mutation, and death.